Written Homework 4

Assigned: Fri Oct 27 2017
Due: Wed Nov 8 2017

Instructions:

- The assignment has to be uploaded to blackboard by the due date. NO assignment will be accepted after 11:59pm on that day.

- We expect that you will study with friends and often work out problem solutions together, but you must write up your own solutions, in your own words. Cheating will not be tolerated. Professors, TAs, and peer tutors will be available to answer questions but will not do your homework for you. One of our course goals is to teach you how to think on your own.

- We require that all homework submissions be neat and easily readable. We recommend using a word processor like Microsoft Word or LaTeX for your submissions. If you scan your homework, you may lose points if the scan is not legible.

- To get full credit, show INTERMEDIATE steps leading to your answers, throughout.

- You can give answers that are probabilities as either fully simplified fractions or decimal answers to at least two nonzero digits.

Problem 1 [22 pts (4,4,6,4,4)]: Black and White

An urn contains 10 balls: 6 white balls numbered 1 through 6, and 4 black balls numbered 1 through 4. We are simultaneously and randomly drawing 2 balls out of the 10.

1. Find the probability of event $A$: the two balls are white.
2. Find the probability of event $B$: the two balls are odd.
3. Are events $A$ and $B$ independent? Prove your answer.
4. Let $X$ be the random variable whose value is the number of white balls in the drawing.
   (a) Write the probability distribution of $X$ in form of a table:

<table>
<thead>
<tr>
<th>Pr[$X=0$]</th>
<th>Pr[$X=1$]</th>
<th>Pr[$X=2$]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

   (b) Find the expected value $E[X]$.

Problem 2 [24 pts (4,4,4,5,7)] Random Stocks

The company EquiCola’s stock price fluctuates in the following way from one day to the next. With probability 0.5, the price remains the same the next day. With probability 0.3, the price increases by one dollar. With probability 0.2, the price decreases by one dollar.
1. What is the expected change in value for the stock from one day to the next?

2. What is the expected change in value for the stock over the course of a week (7 days)? What principle allows you to make this conclusion?

3. What is the probability that the stock does not have a single decline over the course of the whole week? (To two decimal places.)

4. What is the probability that the stock declines no more than once over the course of the week? (To two decimal places.)

5. What is the probability that the price changes every day for the full week, and ends a dollar higher than it started? (To two significant digits.)

Problem 3 [18 pts (2,4,4,4,4)]: Texas Hold 'Em

Note: For this problem, be careful about taking into account cards that you know aren’t in the deck or in the opponents’ hands.

You have just sat down to a game of Texas Hold 'Em, a poker variant played with a normal 52 card deck (2-10, Jack, Queen, King, Ace, each in 4 suits: Clubs, Spades, Hearts, Diamonds). The game has the following steps.

1. Each player is dealt two cards face down.
2. There is a round of betting. Players who don’t want to bet can “fold” and bow out without revealing cards.
3. Three cards are revealed simultaneously in the center (the “flop”).
4. There is another round of betting.
5. Another card is revealed in the center (the “turn”), and there is another round of betting where players may fold.
6. A final card is revealed in the center, and there is a final round of betting. Players reveal their hands, and whoever can make the best 5-card poker hand, using some combination of their 2 cards and the 5 shared cards, wins the pot of money. (You can assume in the problems that follow that we will tell you the relevant facts about values of poker hands.)

i. Calculate the number of possible two-card poker hands you could be dealt.

ii. It turns out you are dealt two Jacks – the black ones (Clubs and Spades). Calculate (a) the number of possible two-card poker hands you could be dealt that are pairs of Jacks, Queens, Kings, or Aces, (b) the probability of being dealt such a pair, to two significant figures.

iii. Two other players fold, leaving just one other player, and the flop is revealed: it’s Ten of Hearts, Ten of Diamonds, Ten of Clubs. You now have a “full house” – two of one card, three of another – which is a pretty good hand. But given that the other player is holding neither your two cards nor any of these three, what is the likelihood the other player is holding the last ten, and therefore beating you with a four-of-a-kind? (To three significant digits.)

iv. What is the probability now that the “turn” and the “river” will both be Jacks? (To three significant digits.)
v. The “turn” comes, and it’s the Jack of Hearts. That turned out well for you. But what is
the probability now that your opponent is holding the final ten? (To three significant digits.)

In the final showdown, you each show your cards. Your opponent wins with the Ace of Diamonds
and the Ten of Spades. You resist the urge to blurt, “What are the chances?” – because you know.

Problem 4 [18 pts (5,4,4,5)]: At the Airport

1. Suppose a bomb detector at the airport has a 95% chance of detecting a bomb if there is one
   in a piece of luggage, but it has a 1% chance of falsely detecting a bomb within an arbitrary
   innocent piece of luggage. Suppose only 1 in 100,000 pieces of luggage actually contains a
   bomb. Calculate the conditional probability of a piece of luggage containing a bomb, given
   that the detector is claiming there is such a bomb inside.

2. What is the probability that three bomb detections are all false alarms?

3. How many detection events must occur until there is one real bomb among them in expecta-
   tion?

4. Now suppose there is a liquid detector with a 95% chance of detecting liquid if luggage
   contains some, and a 1% chance of detecting liquid if there is none. 1 in 5 pieces of luggage
   actually contains a liquid. What is the probability that a piece of luggage contains a liquid,
   given that the detector claims there is liquid?

Problem 5 [18 pts (6, 12)]: Bags of Words

1. Make an intuitive prediction as to which of the words “entropy” and “bookkeeper” has the
   higher entropy (treating the letters as symbols), and explain your prediction. Then calculate
   the entropies of the two words and determine whether you were right. (For each word, treat
   its own symbols as the only symbols that exist.)

2. A Markov babbler is a program trained to generate random sequences of words that match the
   transition probabilities of some target text it was trained on. These transition probabilities
   are conditional probabilities of each possible next word given the previous word, and they can
   be calculated by counting how often each word follows each other one. For example, if the
   training text were TO BE OR NOT TO BE THAT BE THE QUESTION, P(“BE” | “TO”)
   = 1 since BE is the only word that ever follows TO in the training text, but P(“OR” | “BE”)
   = 1/3 since three different words can follow “BE.” The babbler might then end up saying
   something like TO BE THAT BE OR NOT TO BE THAT BE THE QUESTION. (It would
   probably be more interesting with more training text.) It’s called a Markov babbler because
   it effectively samples a Markov chain and babbles the word corresponding to each new state
   as it enters it.

   Consider the babbler trained on the short text LOGIC IS THE PLAN THE PLAN IS MATH.
   Assume it wraps around to the beginning so that “LOGIC” is considered to follow “MATH.”

   (a) Construct the transition matrix that corresponds to the Markov chain that describes
   this babbler’s behavior. Please assume the order of the columns and rows is LOGIC, IS,
   THE, PLAN, MATH.
(b) Use the iterative method described in lecture and the textbook to find the frequency of each word in the babbler’s output as it runs forever. If you use code, please include it in your submission.

(c) Guess the simple fractions your decimal numbers are approximating, and verify that this is the correct stationary distribution by substituting into appropriate equations derived from the transition matrix.